A GENERALIZATION OF A TEST OF COPLANARITY BASED ON THE FISHER DISTRIBUTION

BY

JOHN HUBERT

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE

OF MASTER OF SCIENCE

DEPARTMENT OF MATHEMATICS

UNIVERSITY OF ALBERTA

EDMONTON, ALBERTA

APRIL, 1967

UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies for acceptance, a thesis entitled "A GENERALIZATION OF A TEST OF COPLANARITY BASED ON THE FISHER DISTRIBUTION", submitted by JOHN HUBERT in partial fulfilment of the requirements for the degree of Master of Science.

ABSTRACT

Fisher (1953) derived a distribution of directions where the field of possible observations was the surface of the unit sphere. The distribution of the elementary errors over this surface had a frequency density proportional to $e^{\kappa \cos \theta}$ where θ is the angle between an observed direction and a 'true' direction, where $\theta = 0$ the density is a maximum and where κ is a parameter. Watson (1960) derived a test of whether the mean directions of a set of populations each distributed according to the Fisher distribution all lie in the same plane normal (i.e., $\alpha = \pi/2$) to a prescribed direction. In the present thesis, a generalization of this test of coplanarity to any angle α is derived. A thorough account of the history, development and application of the Fisher distribution is given. A summary of other significance tests based on this distribution is also provided.

ACKNOWLEDGEMENTS

I am indebted to Dr. John R. McGregor for suggesting to me the subject and for patiently assisting me during the research and preparation of this thesis.

TABLE OF CONTENTS

CHAPTER I

INTRODUCTION

The spherical normal distribution on the unit sphere was derived in 1953 by R. A. Fisher [27] in connection with the direction of remanent magnetization of specimens from the same source of rock. In the present thesis this distribution is called the 'Fisher distribution'. Fisher has also derived most of the basic distribution theory. The specific test of whether the mean directions of a set of populations, distributed according to the Fisher distribution, all lie in the same plane normal to a prescribed direction was derived in 1960 by G. S. Watson [96]. This significance test is called the 'test of coplanarity' and this thesis is primarily concerned with the generalization of this test and reviewing the history of the theory associated with it.

In Chapter II a comprehensive review of the distributions associated with measurements of directions is given, and their relationships with the Fisher distribution are demonstrated. Also the Gaussian normal distribution, the von Mises distribution and the Fisher distribution are derived in detail. Many references of their applicability are given.

A variety of significance tests of hypotheses based on the Fisher distribution are discussed in Chapter III. The necessary density functions of the required statistics for these tests are also provided.

The Watson test of coplanarity for the sphere and its generalization are derived in detail in Chapter IV.

A list of references completes the thesis.

CHAPTER II

A REVIEW OF THE EARLIER WORK IN THE DISTRIBUTIONS CONCERNING MEASUREMENTS OF DIRECTIONS

Gauss [29] and Laplace [47] considered measurements of the direction of celestial bodies with respect to the earth. Because of inescapable errors of observations, repeated measurements of the same direction did not coincide but rather tended to deviate about a central location, generally in a very small radius. To cope with this variation, Gauss developed what is known today as the Theory of Errors.

The Gaussian theory assumes that the measurements may be approximated linearly. In recent years there appear physical quantities which cannot be usefully approximated linearly and in fact only the orientation of the measured quantities or variables is of interest. These variables may be regarded as vectors from an origin, ending in points on the circumference of a sphere, where, since the length of the vector is unimportant, the radius is taken as 1. Then the variables of importance are the usual polar coordinates θ and ϕ for the sphere and θ for the circle. The statistical examination of such variables is known as the statistics of directions.

The distributions involving directions can be divided into four basic categories: the uniform distribution, the Brownian motion distribution, the von Mises distribution, and the Fisher distribution. This

chapter will review their development and show how these four distributions are interrelated.

2.1 The Uniform Distribution.

The most extensively studied distribution of directions is undoubtedly the uniform distribution, known also as the isotropic distribution by physicists and astronomers. The first consideration of this distribution was in 1734 by Daniel Bernoulli [4] who questioned whether the close coincidence of the orbital planes of the then known six planets could have arisen by chance. Since each orbit determines a directed line, namely the line normal to the plane and directed in the sense that the motion of the planet is counter-clockwise around it, each orbital plane corresponds to a point on the sphere in the following sense: since a line through a point meets the unit sphere centred there at exactly two points and if it is a directed line these points can be distinguished, then there is a one-to-one mapping of the set of directed axes onto the unit sphere. Hence a distribution defined on the surface of the sphere defines a distribution of directed lines in space and conversely.

This problem was interpreted as a problem of testing the agreement with the uniform distribution on the sphere, and, in fact, Bernoulli applied a test of significance which rejected the uniform distribution on the evidence of observed orbits.

The uniform distribution occurs next in the literature with

apparently two different problems, but as will be seen later, the problems are very similar. In 1905, Karl Pearson [58] proposed the following problem:

> "A man starts from a point O and walks ℓ yards in a straight line; he then turns through any angle whatever, and walks another ℓ yards in a second straight line. He repeats this process n times. I require the probability that after these n stretches he is at a distance between r and r + dr from his starting point O ."

The theory of 'random walk' (see Spitzer [80]) and the diversified problem of 'random flights' (see Rayleigh [69]) have their origin from this inquiry.

Essentially, Pearson was assuming a sample of n random angles α_i $(i = 1,2,\ldots,n)$ from the uniform distribution

$$f(\alpha) = \frac{1}{2\pi} , \tag{2.1.1}$$

and his required probability is

$$dP_n(r,\ell) \tag{2.1.2}$$

where $P_n(r,\ell)$ denotes the probability that the man is at most a distance r after n stretches of ℓ yards each from the point O .

If we define the probability density function $f_n(r,\ell)$ by the relation

$$f_n(r,\ell) = \frac{d}{dr} P_n(r,\ell)$$

or equivalently

$$P_n(r,\ell) = \int_0^r f_n(x,\ell)dx \quad ,$$

then

$$f_n(r,\ell)dr = \frac{d}{dr} P_n(r,\ell)dr$$

represents the probability that the distance between O and the man is between r and $r + dr$.

Within a year three solutions were proposed. The first was by Rayleigh [68] who stated that

> "This problem ... is the same as that of the composition of n iso-periodic vibrations of unit amplitude and of phases distributed at random ... ."

He actually considered this problem in 1880 (see [67]). Rayleigh provided only an asymptotic approximation to the solution for large n in two dimensions. His solution was

$$f_\infty(r,\ell) = \frac{2r}{n\ell^2} e^{-\left(\frac{r^2}{n\ell^2}\right)} \quad . \tag{2.1.3}$$

Pearson [59] acknowledged this solution to be correct only when n is infinite (hence the subscript ∞), and mentioned that it gives a close approximation for large, finite values of n over the range $0 \leq r \leq n\ell$. (The frequency must necessarily be zero for values of r greater than $n\ell$.)

The second solution was given by Pearson [60] himself. (His

results were published in a rather inaccessible journal and we state his solution in a form given in Horner [38] and in Durand and Greenwood [17].) Pearson's solution was

$$(2.1.4) \qquad f_n(r,1) = \frac{2re^{-\frac{r^2}{n}}}{n} \sum_{i=o}^{\infty} c_i L_i\left(\frac{r^2}{n}\right)$$

where the $L_i(\frac{r^2}{n})$ are Laguerre polynomials (as defined in Watson [93]) and the c_i are as follows:

$$c_o = 1$$

$$c_1 = 0$$

$$2!\, c_2 = \frac{-1}{2n}$$

$$3!\, c_3 = \frac{-2}{3n^2}$$

$$4!\, c_4 = \frac{6n - 11}{8n^3}$$

$$5!\, c_5 = \frac{50n - 57}{15n^4}$$

$$6!\, c_6 = \frac{-(1892 - 2125n + 270n^2)}{144n^5}$$

(2.1.4) is an infinite series from which the probability density function $f_n(r,1)$ could be calculated. The accuracy can be found by taking a finite number of terms of this series representation and this problem has been analyzed extensively by Lord [54], Greenwood and Durand [31], Durand and Greenwood [17], Esseen [22], and Horner [38].

The third solution was given concisely by Kluyver [45] who

stated:

"I find that the general solution of this problem depends upon the theory of Bessel's function ...".

Essentially, Kluyver showed that

$$P_n(r,\ell) = r \int_0^{\infty} [J_o(\ell x)]^n J_1(rx)dx \quad , \tag{2.1.5}$$

where $J_i(x)$ denotes the Bessel function as defined in Watson [93].

Since

$$\frac{d}{dr}[rJ_1(rx)] = rxJ_o(rx)$$

then

$$f_n(r,\ell) = r \int_0^{\infty} [J_o(\ell x)]^n xJ_o(rx)dx \quad .$$

Kluyver also showed that

$$P_n(1,1) = \frac{1}{n+1} \quad .$$

This integral solution (2.1.5) for the two-dimensional case was subsequently extended to three dimensions in 1919 by Rayleigh [69].

He also derived the result

$$(2.1.6) \qquad P_n(r,\ell) = \frac{2}{\pi}\int_0^\infty \frac{\sin rx - rx\cos rx}{x}\left(\frac{\sin \ell x}{\ell x}\right)^n dx \;.$$

Moreover, in a subsequent paper, Rayleigh [70] derived asymptotic expansions for the linear and circular cases of this problem.

These solutions derived by Rayleigh in 1880 and 1919 were generated by investigations of problems in the theory of sound. With the advancement of telecommunication theory in the 1940's further applications of Rayleigh's solution appeared. In particular, Slack [77] applied Rayleigh's solution to derive some results for problems involved in the theory of multi-channel transmission for telephone systems. Horner [38] applied Rayleigh's solution to theoretical investigations for site error problems in direction finding. Chandrasekhar [10] gave a thorough survey of how Rayleigh's solution can be applied to many complicated problems in physics and astronomy.

The three basic solutions given by Rayleigh, Pearson and Kluyver all deal with the distribution of the length r, where $0 \leq r \leq \infty$, of the resultant of unit vectors whose directions are uniformly distributed either in one or two dimensions and all are approximate in form. The first to derive an exact form for the distribution of r in three dimensions was Fisher [27] in 1953. This form of the distribution will be discussed separately in section 2.4 and is called the Fisher distribution.

In two dimensions Lévy [51] showed that the addition of random variables on the circle can lead to the uniform distribution. He also stated that this holds for three dimensions but only for the special case when the variables are distributed with axial symmetry. Perrin [61] showed that for this special case this result is a direct consequence of the convolution property of the Legendre polynomial series (see Breitenberger [6] and Roberts and Ursell [71]). Breitenberger [7] states that at present no analytical investigations have been made for this case.

In recent years there have appeared many papers on the uniform distribution on the circle and on the sphere. Percentage points, tables and their application to statistics connected with the uniform distribution have already been published (see Stephens [89]). A thorough review of this distribution can be found in a series of papers by Durand and Greenwood [16], [17], [18], [30], [31].

2.2 The Brownian Motion Distribution

Another type of distribution of direction defined on the circle or the sphere which has received extensive study throughout the years is the 'heat flow distribution' or 'diffusion with a point source' or the 'Brownian motion distribution'. It concerns the addition of infinitely many variables with infinitesimal range. The straight-line case was first encountered in 1900 by Bachelier [2]. The Brownian motion distribution on the circle, first derived by deHaas-Lorentz [34] in 1913, leads to the 'wrapped-up normal distribution'. Basically it

is obtained by identifying points on the line which are equivalent mod 2π and summing the normal density over these points. The density function can be expressed as

$$(2.2.1) \quad f(\theta) = \frac{1}{(2\pi\sigma^2)^{1/2}} \sum_{j=-\infty}^{+\infty} \exp\left[\frac{-(\theta-\theta_0-2j\pi)^2}{2\sigma^2}\right], \quad 0 \leq \theta < 2\pi$$

(Gumbel, Greenwood and Durand [33]). Bingham [5] stated that (2.2.1) could be written in the form

$$f(\theta) = \frac{1}{2\pi} \vartheta_3 \left((\theta-\theta_0), e^{\frac{-\sigma^2}{2}}\right),$$

where ϑ_3 is a Jacobi theta function (see Erdélyi [21]). The density (2.2.1) was later studied by many authors: Pólya [64]; Perrin [61], [62]; Wintner [104], [106]; Lévy [49], [50], [51]; Marcinkiewicz [55]; Zernike, [107]; Sommerfeld [78]. It has been shown by Stephens [84] that (2.2.1) can be closely approximated by the von Mises distribution which will be discussed later. Hence it can be expected that the theory and methods for the von Mises distribution will be approximately correct for the Brownian motion distribution on the circle.

On the sphere, the Brownian motion distribution was first studied by Einstein [19] in 1905. It is the distribution, after a fixed length of time, of the position of a particle starting from a fixed point and moving with infinitesimal randomly oriented steps. The density function can be expressed as

$$(2.2.2) \qquad \frac{1}{4\pi} \sum_{n=o}^{\infty} (4n+1) P_{2n}(\cos\theta)\, e^{-2n(2n+1)\sigma^2}, \quad \sigma^2 > 0 ,$$

where θ is the spherical distance from a fixed direction (the starting point of the moving particle), and P_j $(j = 0,1,\ldots)$ are the Legendre [48] polynomials (see Hobson [37]). In 1960 Roberts and Ursell [71] showed that (2.2.2) could be approximated by the Fisher distribution for a suitable choice of parameters.

In 1921 Wiener [103] used a different method to derive this distribution. He used functional analysis to derive this distribution and to prove that it was approximately normal. The proof is based on the theory of convolutions and the geometrical argument that the equi-distribution on the surface of a n-dimensional sphere can be projected orthogonally on a diameter of this sphere. This approach was used by Perrin [61] in 1928 for the case of the circle and was followed in a general and precise manner for the sphere by Lévy [50] in 1938. Many followers extended Lévy's basic work. Some of these are: Schoenberg [75]; Kac [42]; Hartman and Wintner [36]; Wintner [106]; Furry [28]; Favro [24]. The latter two stressed mostly the formal aspects at the expense of further distribution theory. (For further comments, see Chandrasekhar [10], Coulson [11], Quenouille [66], Breitenberger [7], Stephens [81], and Bingham [5]. In Chapter 2 of [10] there is a thorough review of the development of the Brownian motion theory from 1905 to 1943, and his bibliography should be referred to for further references).

Today the Theory of Brownian Motion is an extensive field in statistical mechanics and the number of books and articles on it is immense.

2.3 The von Mises Distribution.

As remarked earlier, it was in 1809 that Gauss developed the theory of errors. His development was in relation to the infinite linear continuum and the actual topological framework (for example, the surface of the sphere) of such measurements is ignored with a certain gain in simplicity. In 1918 von Mises [56] applied the method of Gauss to a circular variate and derived the distribution known as the 'circular normal distribution on the circle' or 'the von Mises distribution'. In 1953, R. A. Fisher [27] considered how the theory of errors would have had to be developed if the observations had in fact involved errors so large that the actual topology had to be taken into account. In fact he stated:

> "If astronomy had involved measurements of direction of such poor accuracy that they were scattered over a large part of a celestial sphere, the distribution theory of directions would have followed different lines."

We shall discuss the Fisher distribution in the next section where it will be evident that the form of the von Mises distribution suggested to Fisher his original density for the sphere.

Before discussing the von Mises distribution a discussion of the Gaussian normal law of errors and the circular variate will be given.

The former is reviewed here because the von Mises derivation follows the same methods of approach as Gauss' derivation. The latter is discussed because it is to the circular variate, not the linear variate, that von Mises applied Gauss' methods.

2.3.1 The Gaussian Normal Law of Errors.

The normal law of errors was developed in 1809 by Gauss [29]. It states that:

> "... the distribution of the measures about the true value is a [linear] normal frequency distribution."

(See Whittaker and Robinson [102], page 220.) Gauss used the well-known maximum likelihood principle of Legendre [48] and the single assumption called the 'postulate of the arithmetic mean':

> "... when any number of equally likely good direct observations of an unknown magnitude are given, the most probable value is their arithmetic mean."

(See Whittaker and Robinson [102], page 215.) The derivation is essentially as follows: Suppose that for the measure of an observed quantity the probability of an error between Δ and $\Delta + d\Delta$ is $f(\Delta)d\Delta$ so that $f(\Delta)$ is the relative frequency of error. If ϵ denotes the least quantity to which the measuring instrument is sensitive, we can suppose that the possible values of any measure proceed by steps of amount ϵ and the probability of an error Δ may be taken as $f(\Delta)\epsilon$.

If $x_1, x_2, \ldots, x_n$ denote n equally likely observations (measurements) of a quantity x whose true value is x_o, then the errors of the observations are $\Delta_i = x_i - x_o$ $(i = 1,2,\ldots,n)$. Since the probability of the error in the ith measurement is $f(x_i - x_o)\epsilon$, then the probability that a set of measures $(x_1, x_2, \ldots, x_n)$ will occur is

$$\epsilon^n \prod_{i=1}^{n} f(x_i - x_o) \quad .$$

If we assume that all values of x are equally likely to be the true value before the observations are made then when the observations have been made the probability that the true value of x lies between x_o and $x_o + dx_o$ is

$$\frac{\prod_{i=1}^{n} f(x_i - x_o)dx_o}{\int_{-\infty}^{+\infty} \prod_{i=1}^{n} f(x_i - x_o)dx_o} \quad .$$

Therefore the 'most probable' value of the true value of x is that x which makes

$$\prod_{i=1}^{n} f(x_i - x)$$

a maximum, or equivalently that x for which

(2.3.1.1) $$\sum_{i=1}^{n} \frac{d}{dx} \ell n\, f(x_i - x) = 0 \quad .$$

But by the assumption that the mean is the most probable value of x,

(2.3.1.1) is equivalent to

$$x = \frac{1}{n} \sum_{i=1}^{n} x_i \quad ,$$

i.e.,

$$(2.3.1.2) \qquad \sum_{i=1}^{n} (x_i - x) = 0 \quad .$$

Since (2.3.1.1) and (2.3.1.2) are equivalent, then for some constant c, we have

$$\frac{d}{dx} \ell n \, f(x_i - x) = c(x_i - x) \quad ,$$

i.e.

$$f(x_i - x) = A \, e^{-\frac{c}{2}(x_i - x)^2} \quad ,$$

where A is a constant. Thus since $\Delta = x_i - x$

$$f(\Delta) = A \, e^{-\frac{c}{2} \Delta^2} \quad .$$

Since the sum of the probabilities of all possible errors is unity,

$$1 = \int_{-\infty}^{\infty} f(\Delta) d\Delta = A \int_{-\infty}^{\infty} e^{-\frac{c}{2} \Delta^2} d\Delta = (\frac{2\pi}{c})^{\frac{1}{2}} A \quad ,$$

i.e.,

$$A = \frac{h}{\sqrt{\pi}} \quad , \qquad h = \sqrt{\frac{c}{2}} \quad .$$

Thus we obtain

$$f(\Delta) = \frac{h}{\sqrt{\pi}} e^{-h^2\Delta^2} , \tag{2.3.1.3}$$

where h is a constant. Equation (2.3.1.3) is a normal density function so that the distribution of measures about the true value is a normal frequency distribution.

2.3.2 The Circular Variate.

As we have remarked before, von Mises [56] applied the above approach of Gauss to a circular variate. The nature of a circular variate can be better understood by the following considerations:

If n events occur during a span of time (e.g. year) and if each event occurs at a certain time, and if each date is considered as a chance variate α then the distribution of events over the time period may be considered a circular distribution of an angle by situating the n events $\alpha_1, \ldots, \alpha_n$ on the circumference of a unit circle so that the α_i are the angles of the radii drawn from the centre. To characterize such circular observations $\alpha_1, \alpha_2, \ldots, \alpha_n$, introduce the rectangular coordinates

$$x_i = \cos \alpha_i$$

$$y_i = \sin \alpha_i ,$$

and let

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$$

define the means. In polar coordinates, let $\bar{a}$ and α_o be the solutions of the equations:

$$\bar{x} = \bar{a} \cos \alpha_o$$

$$\bar{y} = \bar{a} \sin \alpha_o \quad .$$

This solution is unique unless $\bar{x} = \bar{y} = 0$, and is given by

$$\bar{a} = (\bar{x}^2 + \bar{y}^2)^{1/2} \tag{2.3.2.1}$$

$$\alpha_o = \arctan \frac{\bar{y}}{\bar{x}} \quad ,$$

where the quadrant in which α_o lies must be determined by inspection of the signs of $\bar{x}$ and $\bar{y}$. Equation (2.3.2.1) defines the vector length $\bar{a}$ at the mean direction α_o . Clearly

$$\bar{a} = \frac{1}{n} \left[\left(\sum_{i=1}^{n} \sin \alpha_i \right)^2 + \left(\sum_{i=1}^{n} \cos \alpha_i \right)^2 \right]^{1/2} .$$

Since α_o is the mean direction, then α_o must satisfy

$$\sum_{i=1}^{n} \sin (\alpha_i - \alpha_o) = 0 \quad . \tag{2.3.2.2}$$

Then the statistic α_o is the analogue of the average of the linear variate.

2.3.3 Derivation of the von Mises Distribution.

The derivation of the von Mises distribution (or the circular normal distribution) applys the method of Gauss and the maximum likelihood assumption to the deviations of measured atomic weights from integral values ('Ganzzahligkeit'). The derivation is essentially as follows:

If $\alpha_1, \alpha_2, \ldots, \alpha_n$ denote n equally likely observations of a circular variate α whose true value is α_o , then let the 'errors of observations' be

$$\theta_i = \alpha_i - \alpha_o \quad (i = 1,2,\ldots,n) \quad .$$

von Mises asked for a distribution $f(\theta_i)$ such that the likelihood function is a maximum for that α given by

$$\text{(2.3.3.1)} \qquad \sum_{i=1}^{n} \sin (\alpha_i - \alpha) = 0$$

((2.3.3.1) is the condition based on the postulate of the arithmetic mean for circular variables). The maximum of the likelihood function

$$\prod_{i=1}^{n} f\,(\alpha_i - \alpha)$$

occurs when

$$\text{(2.3.3.2)} \qquad \sum_{i=1}^{n} \frac{d}{d\alpha}\, \ell n\, f(\alpha_i - \alpha) = 0 \quad .$$

Since the two sums (2.3.3.1) and (2.3.3.2) are equal for arbitrary values of α_i, then for some constant k

$$\frac{d}{d\alpha} \ln f(\alpha_i - \alpha) = k \sin (\alpha_i - \alpha) \quad . \tag{2.3.3.3}$$

This differential equation has the solution

$$f(\alpha_i - \alpha) = c\, e^{k \cos (\alpha_i - \alpha)}$$

where the two positive parameters c and k satisfy the condition that the integral of the density over the circumference is one, i.e.,

$$\int_0^{2\pi} f(\theta) d\, \theta = 1 \quad .$$

Then

$$c = \frac{1}{\int_0^{2\pi} e^{k \cos \theta} d\theta} = \frac{1}{2\pi\, J_o(ik)} = \frac{1}{2\pi\, I_o(k)} \quad .$$

where $J_o(ik)$ is the Bessel function of the first kind of order zero with imaginary argument ik and $I_o(k)$ is the corresponding modified Bessel function. Thus

$$f(\alpha - \alpha_o) = \frac{e^{k \cos (\alpha - \alpha_o)}}{2\pi\, I_o(k)}$$

where if we let $\alpha_o = 0$ we have

$$f(\alpha) = \frac{e^{k \cos \alpha}}{2\pi\, I_o(k)} \quad . \tag{2.3.3.4}$$

The polar vector of this distribution is usually taken as the $\alpha_o = 0$ line of polar coordinates. The probability can be written in the form

$$P[\alpha < \theta < \alpha + d\alpha] = \frac{e^{k \cos \alpha}}{2\pi\, I_o(k)}\, d\alpha$$

so that this distribution can be considered as the error distribution for the circumference of a circle.

Remarks.

In 1953 Gumbel, Greenwood and Durand [33] christened (2.3.3.4) the density function of the 'circular normal distribution' because it was derived in a way which is strictly analogous to the Gaussian derivation.

Whereas the density (2.3.3.4) is the case $n = 2$ the Fisher distribution (see Section 2.4) as the case $n = 3$. (Recently Bingham [5] derived the analogue of (2.3.3.4) by means of projections for the case of $n \geq 2$.)

We note here that Pólya [65] has proved that this distribution shares some properties analogous to the linear normal distribution. How-

ever, Kac and Kampen [43] as well as Breitenberger [7] have shown that some properties of the normal law cannot be obtained by any but trivial correspondences on the circle.

Some properties of the von Mises distribution are:

i) If $k = 0$, $f(\alpha)$ degenerates to the uniform distribution

$$f(\alpha) = \frac{1}{2\pi} \quad .$$

ii) Since, as k becomes larger, the larger part of the distribution is situated in the neighbourhood of the mode (most probable value) α_o, then the parameter k is a measure of concentration.

iii) If k is large the distribution converges to the linear normal distribution (2.3.1.3). (See Gumbel, Greenwood and Durand [33].)

iv) $1/k$ is analogous to the variance of the linear normal distribution (see Gumbel, Greenwood and Durand [33]).

v) The maximum likelihood estimate of k is given by the solution of

$$I_o'(k) - \bar{a}\, I_o(k) = 0 \quad ,$$

where $\bar{a}$ is defined by (2.3.2.1). (See Arnold [1].)

vi) The maximum likelihood estimate of α_o is given by the solution of

$$\sum_{i=1}^{n} \sin(\alpha_i - \alpha_o) = 0 \quad .$$

(See Greenwood and Durand [31].)

Arnold [1] generalized the circular normal distribution to the bivariate case by considering the centre of gravity of points on a sphere, and also investigated the 'wrapped-up normal' on the surface of a sphere. Breitenberger [7] has also derived the bivariate analogue of the von Mises distribution. Brooks and Carruthers [8] have translated a bivariate normal distribution into polar coordinates with the origin removed from the centre of the distribution and have integrated out with respect to the radius.

Tables, tests and estimation procedures are available in [17], [30], [31], [32], [33], [101]. Recently, Stephens [81], [82], [87], [88] has published a series of papers with accurate tables for the calculation of the maximum likelihood estimate of k and a significance table for the null hypothesis $k = 0$. Much of the basic sampling distribution, however, involves integrals of Bessel functions which have not been tabulated except for the null case $k = 0$. More recently, Downs [14], [15] has given a very exhaustive study of the von Mises distribution.

The problem of generating other circular distributions around the circumference of a circle (e.g., the 'wrapped-up normal' and the Cauchy distribution) have been studied by Perrin [61], Lévy [50], Marcinkiewicz [55], Wintner [105], [106], and Hartman and Wintner [36]. Stephens [84], [85] has shown that on the circle the 'wrapped-up normal'

can be approximated by the von Mises distribution. The 'angular cardioid' distribution as discussed in [33] and the 'covering circle' of a sample from a circular normal distribution appears in Daniels [13]. Goodness-of-fit tests for the circle have been studied thoroughly by Watson [97], [98].

The von Mises distribution has its practical application to geophysical, vital and economic statistics, as shown by Gumbel [32]. Epstein and Sobel [20] have used a variation of this distribution in 'life testing'. Other applications can be found in [15], [30], [31], [78], [101].

2.4 The Fisher Distribution.

In 1953 Fisher [27] derived a distribution of directions known today as the 'spherical normal distribution', the 'von Mises distribution on the sphere' and simply as the 'Fisher distribution'. Before Fisher's derivation, as pointed out in Section 2.1 most measurements in astronomy had involved directions which could be approximated linearly. However, in other sciences such as geology, biology and meteorology there occur measurements of direction which cannot be usefully linearly approximated, as demonstrated by Fairbairn [23], Pincus [63], Runcorn [74], Tucker [91] and Waterman [92] - thus the need for Fisher's distribution.

2.4.1 Derivation.

In his paper, Fisher [27] considered the field of possible

observations to be the surface of a unit sphere and let the distribution of elementary errors over this surface have a frequency density proportional to

$$e^{\kappa \cos \theta}$$

where θ is the angular displacement between an observed vector and a fixed polar vector at which $\theta = 0$ the density is a maximum, provided κ is a positive constant. If θ and ϕ are the usual spherical polar coordinates (so that $\theta = 0$ is the 'North pole' axis), then

$$P[\theta < \theta_o < \theta + d\theta,\ \phi < \phi_o < \phi + d\phi] = f(\theta,\phi)d\theta\ d\phi$$

where $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. Since the surface area within limits $d\theta$ and $d\phi$ is proportional to $\sin \theta\ d\theta\ d\phi$ then

$$f(\theta,\phi)d\theta\ d\phi = C\ e^{\kappa \cos \theta} \sin \theta\ d\theta\ d\phi$$

where C is the constant of proportionality. Since we require

$$\int_{\theta = 0}^{\pi} \int_{\phi = 0}^{2\pi} f(\theta,\phi)d\theta\ d\phi = 1 ,$$

then

$$\int_{\theta = 0}^{\pi} \int_{\phi = 0}^{2\pi} (-C\ e^{\kappa \cos \theta})d \cos \theta\ d\phi = 1$$

so that

$$C = \frac{1}{-2\pi \int_{\theta = 0}^{\pi} e^{\kappa \cos \theta} d \cos \theta} = \frac{\kappa}{2\pi\ (e^{+\kappa} - e^{-\kappa})}$$

i.e.,

$$C = \frac{\kappa}{4\pi \sinh \kappa} .$$

Thus

$$(2.4.1) \qquad f(\theta,\phi) = \frac{\kappa \sin\theta\, e^{\kappa\cos\theta}}{4\pi \sinh\kappa} ,$$

where κ is taken positive so that the distribution is located at the (north) pole (if $\kappa < 0$, this gives the same distribution with polar direction reversed and hence need not be considered).

2.4.2 Properties.

If $\kappa = 0$ the distribution is uniform all over the spherical surface. When κ is large the distribution is confined to a small portion of the sphere in the neighbourhood of the maximum (the pole) and the distribution is approximately a linear normal distribution; for, by letting

$$\cos\theta = 1 - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - + \dots$$

then for large κ and small θ, the approximation becomes

$$A\, e^{-\kappa \frac{\theta^2}{2}} ,$$

where A is some constant.

Since large values of κ lead to small dispersions, κ is an accuracy parameter. (For an explanation of the significance of κ in numerical terms see Watson and Irving [100], page 290.)

If we let (ℓ_i, m_i, n_i) be the direction cosines of an observed direction and if (λ, μ, ν) are the direction cosines of the mean direction from the centre of the sphere then

$$\cos\theta = \lambda \ell_i + \mu m_i + \nu n_i \quad .$$

Fisher [27] has shown that for a sample of size N the maximum likelihood estimator of (λ, μ, ν) is the set of direction cosines (ℓ, m, n), say, of the vector resultant of the sample unit vectors, i.e.,

$$\ell = \sum_{i=1}^{N} \ell_i \,/\, R$$

$$m = \sum_{i=1}^{N} m_i \,/\, R$$

$$n = \sum_{i=1}^{N} n_i \,/\, R \quad ,$$

where R denotes the length of the vector resultant; i.e.,

$$R^2 = \left(\sum_{i=1}^{N} \ell_i\right)^2 + \left(\sum_{i=1}^{N} m_i\right)^2 + \left(\sum_{i=1}^{N} n_i\right)^2 \quad .$$

If (λ, μ, ν) is estimated, then the maximum likelihood estimator of κ, $\hat{\kappa}$, say, is the solution of

$$\coth\hat{\kappa} - \frac{1}{\hat{\kappa}} = \frac{R}{N} \quad .$$

When N is large and R is near N, this has the approximate solution

$$\hat{\kappa} = \frac{N}{N - R} \quad .$$

If the pole (λ, μ, ν) is known, and χ denotes the

angle between the resultant and (λ, μ, ν), then the maximum likelihood estimator of κ, $\hat{\kappa}_o$, is the solution of

$$\coth \hat{\kappa}_o - \frac{1}{\hat{\kappa}_o} = \frac{R \cos \chi}{N} .$$

If

$$X = R \cos \chi$$

the approximate estimator of κ is

$$\hat{\kappa}_o = \frac{N}{N - X} .$$

Other properties of the Fisher distribution will be given in Section 3.1 where these properties will be used to describe various tests of significance.

2.4.3 Remarks.

The probability density (2.4.1) on the sphere was first called the density function characterizing the Fisher distribution by Watson [94] in 1956. Although Arnold [1] in 1941 described this three dimensional case before Fisher did in 1953, Arnold's description is from the view of estimation while Fisher's work is a more thorough approach. Arnold's approach was to extend the Pearson [58] 'random walk' problem to the circle and then to the sphere. The problem can now be stated as follows:

"A particle, initially at a given point on a circle
or sphere, moves in an arbitrary direction along the

> circle or along an arbitrary great circle through the initial point on the sphere at a fixed speed for a fixed time. It again chooses an arbitrary direction or arbitrary great circle through the new position along which it moves at the same speed for the same amount of time. This process is repeated many times. We let the time occupied by one unit of this motion approach zero, keeping the total elapsed time for the whole process of the same order of magnitude. We ask for the probability that the particle will be in a given section of the circle or spherical surface after a given total elapsed time for the whole process."

Recently, Stephens [85] studied this random walk on the circumference of a circle, and Roberts and Ursell [71] have given a more general investigation of the random walk on the surface of the sphere.

Breitenberger [7] has demonstrated why the Fisher distribution may be justifiably called the 'spherical normal distribution' by investigating the properties that are preserved under spherical transformations. Also he has extended (2.4.1) to the bivariate analogue. The Fisher distribution has also been derived in several physical theories, see, for example, Jeffreys [40] or Joos [41] who show that the Fisher distribution is a generalization of a two-dimensional form used in mechanics. Lord [52], [53], [54] also has investigated the spherical normal distribution.

In Section 2.2, we have pointed out how the Brownian motion distribution in three dimensions is related to the Fisher distribution. In Section 2.4.2 the relation to the uniform distribution is given.

And, as we have seen in Section 2.3, the von Mises distribution is the two dimensional case of the Fisher distribution. The n-dimentional case has been derived by Stephens [81]. These higher dimensional distributions are shown to be derivable from lower dimensional ones by rotations (see Bingham [5], Downs [14] and Greenwood [30]). Downs [14] and Bingham [5] have demonstrated still more relationships between the von Mises and the Fisher distributions.

Following Fisher's paper there appears a series of papers by Watson [94], [95], [96], [99], Watson and Irving [100], and Watson and Williams [101], in which the theory of this distribution has been further developed with confidence techniques being applied to inference problems (Greenwood [30] gives an excellent review of the distribution theory as well as some other distributions involving angular variables.) In 1960 Roberts and Ursell [71] demonstrated that Perrin's [61] work of 1928 and Fisher's work hardly differ numerically. Recently in a series of papers by Stephens [82], [83], [86], [87], [88] tables, graphs, nomograms were constructed for a variety of significance tests (see Chapter III).

As stated earlier, certain measurements that cannot be usefully approximated linearly in some sciences generated this theory. In fact, Fisher's original paper was concerned with the analysis of measurements of directions of remanent magnetism in lava rocks. Although many applications in other sciences have been found, it is primarily in geology that this theory has been most useful. (See, for example, Fisher [27], Runcorn [74], Pincus [63],

Watson [94], [96], Watson and Irving [100] and Watson and Williams [101].) Some other papers which also use actual geological data to illustrate the use of the Fisher distribution are Krumbein [46], Arnold [1], and Selby [76] in which the problem of determining the preferred orientation (direction) of optical axes of specimens of crystals, rocks and pebbles is analyzed. The folding of a layer of rock has been analyzed by Watson [99] and Bingham [4].

Krumbein [46] has mentioned that these orientation problems are similar to the problem of determining the direction of neutron charges. Breitenberger [7] suggests problems of paramagnetism and of orientational polarizability in mechanics are also related.

Fisher [27] and Krumbein [46] also mention that in astronomy the problem of determining the position of stars is a related application.

In zoology, Watson [98] (as well as Pearson [60]) uses data from the migration of displaced birds such as pigeons as an illustration.

We note here that although the Fisher distribution is easy to work with it has one limitation: it is unapplicable to distributions which are elliptical about a pole. Breitenberger [7] states that this is due to:

> "... the fault that it contains only one parameter. This axial symmetry often precludes application in palaeomagnetic research when markedly elongated samples occur (for striking instances, see Watson and Irving [100])."

Since the practical applications of Fisher's distribution usually involve tests of some kind or another and a generalized test involving the Fisher distribution will be given in Chapter IV, a chapter reviewing the different significance tests will be given next.

There are other distributions involving the sphere with densities similar to that used by Fisher. Some of the distributions are:

i) the 'girdle distribution' of Selby [76] which utilizes densities proportional to

$$e^{-\kappa |\cos \theta|} \quad \text{and} \quad e^{\kappa \sin \theta} .$$

Stephens [88] has analyzed these still further and has published significance tables.

ii) the 'equatorial distribution' on the sphere, which involves densities proportional to

$$e^{\kappa \cos^2 \theta} .$$

This density was also mentioned in several connections by Arnold [1] and Breitenberger [7] and was studied in detail by Watson [99] and recently a definitive discussion of this density have been given by Bingham [5].

iii) the 'bipolar distribution' for which Krumbein [46] (as well as Arnold [1] and Breitenberger [7]) has suggested the densities depend on

$$e^{\kappa \cos 2\theta} \quad \text{or} \quad e^{\kappa \cos^2 \theta} .$$

CHAPTER III

SIGNIFICANCE TESTS FOR THE FISHER DISTRIBUTION ON THE SPHERE

Practical applications of Fisher's distribution require a series of significance tests entirely analogous to those in current use for the linear normal distribution. The tests can be classified in the following two categories:

a) tests of hypotheses concerning κ ,

b) tests of hypotheses concerning the polar vector.

Before stating these tests it will be necessary to have some preliminary distribution theory. The two types of tests are outlined in Sections 3.2 and 3.3.

3.1 Introductory Results.

To the geometry of the sphere, Fisher [27] applied the argument of Irwin [39] and Hall [35] on the rectangular distribution and his own geometrical argument developed earlier (Fisher [26]) to derive the following functions:

$$(3.1.1) \qquad P_N(x) = \sum_{s=0}^{N} \binom{N}{s} (-1)^s < N - x - 2s >^{N-1}$$

$$(3.1.2) \qquad Q_N(x) = \sum_{s=0}^{N} \binom{N}{s} (-1)^s < N - x - 2s >^{N-2} ,$$

where the notation $<x>$ means

$$<x> = \begin{cases} x\ , & \text{if } x \geq 0 \\ 0\ , & \text{if } x \leq 0 \end{cases} \quad .$$

Stephens [81] calls $P_N(x)$ and $Q_N(x)$ the 'Fisher polynomials' of degree $(N-1)$ and $(N-2)$ respectively. Watson [94] introduced the notation $<x>$. Fisher [27] also defines the function

$$(3.1.3) \qquad \phi_N(x) = \frac{1}{(N-2)!}\, Q_N(x) \quad .$$

<u>Case 1</u>: <u>1 SAMPLE, $\kappa \neq 0$</u> .

By means of the polynomials Fisher [27] derived the necessary sample distributions: If N is the number of observations in a sample from a population possessing the Fisher distribution, and if R is the length of the vector resultant of these observations and if Θ is the angle between R and the x-axis (i.e., angular error) and X is the component of R on this axis, then the three statistics which are considered, are R, X and c where $c = \cos\Theta$. The following relations were derived by Fisher [27]:

$$(3.1.4) \qquad X = \sum_{i=1}^{N} x_i = \sum_{i=1}^{N} \cos\theta_i = R\cos\Theta = Rc$$

where the x-axis will be the axis $\theta = 0$ (see also Stephens [81]);

$$(3.1.5) \qquad f_\kappa(R,c) = \left(\frac{\kappa}{2\sinh\kappa}\right)^N e^{\kappa Rc}\, \phi_N(R)R \quad ,$$

the joint density function of R and c ;

$$f_\kappa(X) = \left(\frac{\kappa}{2 \sinh \kappa}\right)^N \frac{e^{\kappa X}}{(N-1)!} P_N(X) \quad , \tag{3.1.6}$$

the density function of $X = Rc$;

$$f_\kappa(R) = \left(\frac{\kappa}{2 \sinh \kappa}\right)^N \frac{2 \sinh \kappa R}{\kappa} \phi_N(R) \quad , \tag{3.1.7}$$

the density function of R ;

$$f_\kappa(R|X) = \frac{\phi_N(R)\ (N-1)!}{P_N(X)} \quad , \tag{3.1.8}$$

the conditional density of R , given X .

Case 2: 1 SAMPLE, $\kappa = 0$.

The case of $\kappa = 0$ is called the case of randomness or uniformity. From Case 1 when $\kappa \neq 0$ the form of the following density functions will be immediately obvious. We have

$$f_0(R) = \frac{R}{2^{N-1}} \phi_N(R) \quad , \tag{3.1.9}$$

the density function or R when $\kappa = 0$. We note here that Rayleigh [69] found

$$f_0(R) = \frac{2}{\pi} R \int_0^\infty \frac{\sin^N x}{x^{N-1}} \sin Rx \, dx \quad . \tag{3.1.10}$$

Stephens [81] has verified that (3.1.9) and (3.1.10) are equivalent.

When N is large, it follows from (3.1.9) that the asymptotic density function of R is

$$(3.1.11) \qquad f_0(R) = \frac{3\sqrt{6}}{\sqrt{\pi}} \frac{R^2}{N^{3/2}} \exp\left[\frac{-3R^2}{2N}\right] .$$

Fisher [27] has derived the following functions when $\kappa = 0$:

$$(3.1.12) \qquad f_0(R,c) = \frac{1}{2^N} \phi_N(R)R \quad ,$$

the joint density function of R and c ;

$$(3.1.13) \qquad f_0(X) = \frac{P_N(X)}{2^N (N-1)!} \quad ,$$

the density function of $X = Rc$.

Case 3: p SAMPLES, $\kappa \neq 0$.

Fisher [27] has generalized the results of Case 1 to any number of samples: Suppose we have p samples each from the same population possessing the Fisher distribution. Let N_i denote the number of observations in the ith sample, R_i denote the length of the vector resultant of the ith sample, R denote the length of the vector sum of the resultants of the separate samples and

$$N = \sum_{i=1}^{p} N_i \quad .$$

Then the joint density of all the R_i and R becomes

$$(3.1.14) \qquad \left(\frac{\kappa}{2 \sinh \kappa}\right)^N \frac{2 \sinh \kappa R}{R} \prod_{i=1}^{p} \phi_{N_i}(R_i) \quad ,$$

the density of R is

$$(3.1.15) \qquad \left(\frac{\kappa}{2 \sinh \kappa}\right)^N \frac{2 \sinh \kappa R}{\kappa} \phi_N(R) \quad ,$$

and the conditional density of all the R_i , given R , is

$$(3.1.16) \qquad \frac{\prod_{i=1}^{p} \phi_{N_i}(R_i)}{\phi_N(R)} \quad .$$

(See also Watson and Williams [101].)

Case 4: 1 SAMPLE, κ LARGE.

The limiting or asymptotic distributions for the case of large κ have been derived by Watson [94]. They are:

$$(3.1.17) \qquad 2\kappa (N-X) \simeq \chi^2_{2N}$$

$$(3.1.18) \qquad 2\kappa (N-R) \simeq \chi^2_{2N-2}$$

$$(3.1.19) \qquad 2\kappa (R-X) \simeq \chi^2_{2}$$

where χ^2_{2N} is the chi-square distribution with 2N degrees of freedom. If R is large so that R is near N , (3.1.18) follows from (3.1.7). Indeed, under this assumption (3.1.7) reduces to

$$\kappa^{N-1}\, e^{-\kappa(N-R)}\; \frac{(N-R)^{N-2}}{(N-2)!} \quad ,$$

and making a change of variable from R to $x = 2\kappa(N-R)$, we obtain

$$\frac{e^{-\frac{x}{2}}\,(\frac{x}{2})^{N-2}}{2(N-2)!} = \frac{e^{-\frac{x}{2}}\,(\frac{x}{2})^{\frac{2(N-1)-2}{2}}}{2\,\Gamma\,(\frac{2(N-1)}{2})} \quad ,$$

which is the density function of a χ^2 variable with $2(N-1)$ degrees of freedom. Thus

$$2\kappa(N-R) \simeq \chi^2_{2(N-1)} \quad .$$

Similarly (3.1.17) can be verified.

We note here for completeness sake that for the two-dimensional case, i.e., for the case of the circle, the corresponding asymptotic distributions are

$$2\ \kappa\ (N-X) \simeq \chi^2_N$$

$$2\ \kappa\ (N-R) \simeq \chi^2_{N-1}$$

$$2\ \kappa\ (R-X) \simeq \chi^2_1$$

and can be found in Watson and Williams [101].

Case 5: 1 SAMPLE, LARGE N .

We require the following approximate forms of the Fisher polynomials $P_N(x)$ and $Q_N(x)$ defined in (3.1.1) and (3.1.2), when

N is large,

$$P_N(R) \simeq \frac{(N-1)!\, 2^{N-1} \sqrt{6}}{\sqrt{\pi}\, \sqrt{N}} \, e^{-\frac{3R^2}{2N}}$$

$$Q_N(R) \simeq \frac{(N-2)!\, 2^{N-1}\, 3\sqrt{6}}{\sqrt{\pi}\, N^{3/2}} \, R\, e^{-\frac{3R^2}{2N}} \quad .$$

Also we have

$$\frac{d}{dx} P_N(x) = -(N-1)\, Q_N(x)$$

$$\phi_N(R) \simeq \frac{2^{N-1}\, 3\sqrt{6}}{\sqrt{\pi}\, N^{3/2}} \, R\, e^{-\frac{3R^2}{2N}} \quad .$$

(See Stephens [81].) From Case 1, i.e., from (3.1.6) and (3.1.7), the approximation for the density function of R becomes

$$f_\kappa(R) \simeq \left(\frac{\kappa}{\sinh \kappa}\right)^N \frac{3\sqrt{6}}{\kappa\sqrt{\pi}\, N^{3/2}} \, R \sinh(\kappa R)\, e^{-\frac{3R^2}{2N}} \quad ,$$

and the density function of X becomes

$$f_\kappa(X) \quad \left(\frac{\kappa}{\sinh \kappa}\right)^N \frac{\sqrt{6}}{2\sqrt{\pi}\, \sqrt{N}} \, e^{\kappa X - \frac{3X^2}{2N}} \quad ,$$

while the asymptotic distribution for large N and $\kappa = 0$ is

(3.1.20) $$\frac{3}{N} R^2 \simeq \chi_3^2 \quad .$$

This has been concisely proven in Stephens [81] and provides an approximate test for randomness (see Section 3.2.2). Also if $\kappa = 0$ and R is very close to N such that $N - R < 2$, then (3.1.8) becomes

$$f_0(R) \simeq \frac{2R}{2^N\,(N-2)!}\,(N-R)^{N-2} \tag{3.1.21}$$

(see Stephens [81]).

3.2 Test of Hypotheses Concerning κ.

3.2.1 Exact Tests for κ.

Watson [94] has suggested that when the polar vector of the population is known, an exact significance test of κ can be achieved by using (3.1.6) since $X = Rc$ is a sufficient statistic of κ (see proof in Stephens [81] page 78). The construction of confidence intervals for κ when the polar vector is known is given in detail by Stephens [88] by using significance tables.

3.2.2 Tests of the Null Hypothesis $\kappa = 0$.

Since when $\kappa = 0$ the density function of the Fisher distribution is constant, then the null hypothesis $\kappa = 0$ is the hypothesis of randomness (or uniformity). Bruckshaw and Vincenz [9] suggested that the value of the density of R, with $\kappa = 0$, computed for the observed value, be compared with the modal value of the density of R (with $\kappa = 0$). However it is Watson [94] who first provides a statistical test of randomness. It is based on the following argument:

> "Given a sample of size N, the length R will be large if the sample shows preferred direction and small if it does not. Assuming that there is no preferred direction (i.e., $\kappa = 0$) a value R_0, say, may be calculated which will be exceeded by R with any stated probability."

Thus an exact test may be made of the hypothesis that a sample of N vectors is randomly distributed by finding the probability that $R > R_0$:

$$P[R > R_0] = \int_{R_0}^{N} f_0(R)\, dR = \alpha \ ,$$

where, for the Fisher distribution, α may be explicitly calculated by using the Fisher form of $f_0(R)$ given in (3.1.9).

In a subsequent paper Watson [95] gave this test and included significance points of R_0 for various probabilities and sample sizes: $\alpha = 5\%$ and $\alpha = 1\%$, $N = 5$ to $N = 20$. Stephens [81] recomputed this table and extended α to four values: 1%, 2%, 5%, and 10% for $N = 4$ to $N = 20$. (See his Table 2.2.) Thus to carry out the test it is merely necessary to enter Stephens' table at the row corresponding to the number of observations in the sample in order to find the value of R_0 which will be exceeded with a given probability in sampling from a population in which $\kappa = 0$. Stephens [86] recently has extended his table to include values of $N = 4$ to $N = 25$. He describes the exact test for randomness as follows:

i) Find R.

ii) Find R_0 in the table for appropriate N and α .

iii) If $R > R_0$, reject the hypothesis $\kappa = 0$ at the α-significance level.

When κ is large there also exists an approximate test for randomness. It was first given by Watson [95] by using the form

$$R^2 \simeq \frac{N}{3} \chi_3^2$$

(see Equation (3.1.20)). Stephens [81], [82], has established this same result but his derivation uses the multi-dimensional form of the Central Limit theorem (see Cramér [12]). The approximate test for randomness for large N , given by Stephens [86] is:

i) Find R .

ii) Find R_0 from the equation

$$R_0^2 = \frac{N}{3} \chi_3^2 (\alpha)$$

where $\chi_3^2 (\alpha)$ is the α-significance level of the chi-square distribution with 3 degrees of freedom, upper tail.

iii) If $R > R_0$, reject the hypothesis $\kappa = 0$ at the α-significance level.

3.2.3 Tests of $\kappa = \kappa_0$.

Watson [94] has suggested that an exact significance test of the null hypothesis $\kappa = \kappa_0$ may be made by using the density of R in

(3.1.7) with $\kappa = \kappa_0$ when the alternative hypothesis is $\kappa > \kappa_0$.

As an approximate test, when κ is large, we use (3.1.17) and (3.1.18), that is,

$$2\,\kappa\,(N-X) \simeq \chi^2_{2N}$$

$$2\,\kappa\,(N-R) \simeq \chi^2_{2(N-1)} \quad .$$

Hence an approximate test of the hypothesis $\kappa = \kappa_0$ can be made by referring $2\kappa_0(N-X)$ or $2\kappa_0(N-R)$ to χ^2-tables with $2N$ and $2(N-1)$ degrees of freedom respectively.

Stephens [87] has constructed significance tables for the exact and approximate tests of $\kappa = \kappa_0$ both when the polar vector is known and unknown.

The derivation of the exact tests in higher dimensions for this case can be found in Stephens [81].

3.2.4 Tests of Several κ's .

Watson [94] first suggested that for a comparison of two κ's an F-test may be used because $1/\kappa$ corresponds to the variance of the distribution. Watson and Irving [100] state that if samples of N_1 and N_2 observations give dispersion estimates k_1 and k_2 then

$$(3.2.4.1) \qquad \frac{k_1}{k_2} = \frac{\text{variance with } 2(N_2 - 1) \text{ degrees of freedom}}{\text{variance with } 2(N_1 - 1) \text{ degrees of freedom}} \quad ,$$

assuming the two populations have the same value of κ . Thus this assumption may be tested since the right hand side of (3.2.4.1) has a F-distribution and values of

$$F = \frac{k_1}{k_2}$$

far from unity suggest that $\kappa_1 \neq \kappa_2$.

For the general case, Watson and Irving [100] have suggested that for several populations the ratio of the largest to the smallest estimates may be used to test the hypothesis that κ is constant over the populations. Watson and Williams [101] suggest that since the test of homogeneity of the values of κ for several results requires the sample resultants, then (3.1.16) would be the basis for any exact significance test for this case. Also, they state that for the case of known polar vectors, there is no practical interest and they do not pursue this case further. For the case of unknown polar vectors they do suggest that (3.1.16), being independent of κ , could be used to construct an exact test not depending on the nuisance parameter κ . However, when both κ and N_i are small they state that nothing is known of suitable tests.

3.3 Tests of Hypotheses Concerning Polar Vectors.

3.3.1 A Single Prescribed Polar Vector.

Fisher [27] has shown (see Watson and Williams [101] for a numerical demonstration) that a test of a prescribed

polar vector may be made by using

$$1 - c \quad \theta = \frac{N - R}{R}\left[\left(\frac{1}{P}\right)^{\frac{1}{N-1}} - 1\right] , \qquad (3.3.1.1)$$

where P is the probability that the cosine of the angle between the resultant and the polar vector is less than c. Thus for example, for a given sample, R can be calculated and for a radius of confidence at the 5% level of confidence we use (3.3.1.1) where N and R are now known and $1/P = 20$ so that θ is determinable. This value gives a cone of confidence for the polar vector of the population from the sample. This test is analogous to the 'Student' [90] t-test for a single sample (see [27] and [94]).

When κ is large and $N - R \leq 2$, Fisher [27] has shown by example that this method can be used as a fiducial test of the hypothesis that the polar vector of the population from which the sample is drawn has a prescribed direction.

Watson and Irving [100] give the approximation of (3.3.1.1) as

$$1 - c = - \frac{\ln P}{k\,N} ,$$

where $c = \cos\theta$ and k is the best estimate of κ and is given by

$$k = \frac{N - 1}{N - R} .$$

Watson [94] suggests that, for the case of a given polar vector

prescribed by hypothesis, the test statistic

$$(N-1)\,\frac{R(1-c)}{N-R} \simeq F_{2,2(N-1)}$$

should be used.

Watson and Williams [101] have developed tests which do not depend on κ : If κ is known, a test of a prescribed polar vector may be made by (3.1.6) and when κ is unknown the test is the analogue of the single sample 'Student' t-test and the density of R given $X = Rc$ is considered.

Stephens [83] reviewed these previous results and provides, by means of nomograms, an exact test for the null hypothesis that a given vector is the polar vector of the distribution when κ is unknown. He discusses two approximate tests for this hypothesis as well as an approximate test that the polar vector lies in a given plane.

Stephens [88] has also constructed nomograms for the case when κ is unknown for exact and approximate tests of the hypothesis that a given vector is the polar vector.

3.3.2 Comparison of Two Polar Vectors.

Watson [94] considered samples of size N_1 and N_2 drawn from two populations and assumed that both populations have equal values of κ and then suggested that the statistic

$$(N-2)\,\frac{(R_1+R_2-R)}{(N-R_1-R_2)} \simeq F_{2,2(N-2)}$$

provided a significance test of the hypothesis that the polar directions of the populations were identical (because if the mean vectors are very different, $R_1 + R_2$ will be much greater than R so that large F indicates significance).

3.3.3 Test for Several Polar Vectors.

To test the equality of several polar vectors for different samples, assuming that all populations have the same κ , Watson [94] suggested that this test would be similar to that of Section 3.3.2 above. The generalization to p populations has been derived by Watson and Irving [100]. They suggest the following test statistic:

$$\frac{2\left(\sum_{i=1}^{p} N_i - p\right)}{2(p-1)} \cdot \frac{\sum_{i=1}^{p} R_i - R}{\sum_{i=1}^{p} N_i - \sum_{i=1}^{p} R_i} \simeq F_{2(p-1),2\left(\sum_{i=1}^{p} N_i - p\right)} ,$$

where the sample from the ith population is of size N_i and has resultant length R_i and R is the length of the vector sum of the resultants of the separate samples. Here, large values of F would suggest that the assumption of identical polar vectors is false because the algebraic sum of the sample resultants

$$\sum_{i=1}^{p} R_i$$

will then be much greater than the length of their vector sum R .

This same generalization can be found in Watson and Williams [101]. Moreover, they derive a test which does not include the nuisance parameter κ . They use the fact that since (3.1.16) is free of κ it provides a possible exact test.

3.3.4 Tests for Coplanarity.

The problem of testing whether three polar vectors are coplanar, that is, that the ends of the three vectors all lie on the same great circle, was first suggested by Watson [94]. He remarks that the method used in the generalization of Section 3.3.3 above would apply but

> "... the estimation equations are no longer directly soluble".

In a subsequent paper, Watson [96] derived a test of coplanarity for p populations by another method, that of likelihood-ratio methods. Because a generalization of this test will be given in Section 4.2 using the same method but slightly different notation, a brief outline of Watson's test of coplanarity is given in the next chapter.

CHAPTER IV

THE TEST OF COPLANARITY

The problem of testing whether a set of vectors all lie in the same plane normal to a given direction is a problem of deriving a test statistic for this case of coplanarity. Watson [96] solved this problem by considering several populations each distributed according to the Fisher distribution on the unit sphere. In this chapter Watson's test of coplanarity is rederived. This is done in Section 4.1. The generalization of this test is derived in detail in Section 4.2.

We remark here that if we have several populations (each of which is distributed according to the Fisher distribution) distributed randomly on the surface of a (unit) sphere and if α is the angle between a known direction and the vectors representing the population means, then Watson's case is that of $\alpha = \pi/2$. The generalization extends the test statistic to values of α from 0 to π (i.e., $0 < \alpha < \pi$). Geometrically, the problem is now a problem of testing whether a set of vectors emanating from the centre of a sphere form a right circular cone for which the vertex is the origin of the sphere and the semi-vertical angle is α. The contour on the surface of the sphere will be a circle. For the particular case $\alpha = \pi/2$ the cone becomes a plane through the centre of the sphere and the contour is a great circle.

As remarked earlier many geologists have used this distribution on the sphere as the natural mathematical model for the surface of the

earth.

4.1 Watson's Test of Coplanarity.

Watson [96] considered p populations randomly distributed on the surface of a unit sphere where each population was distributed according to the Fisher distribution. The probability density function for each population can be written in the form

$$\frac{\kappa}{4\pi \sinh \kappa} e^{\kappa \cos \beta_i} \qquad (i = 1,2,\dots,p)$$

where the parameter κ is assumed to be the same for each population,

$$\cos \beta_i = \ell_i \lambda_i + n_i \mu_i + n_i \nu_i \quad ,$$

where (ℓ_i, m_i, n_i) are the direction cosines of an observed direction from the ith population and $(\lambda_i, \mu_i, \nu_i)$ are the direction cosines of the mean vector of the ith population. If a sample of size N_i is taken from the ith population $(i = 1,2,\dots,p)$ and if we let $N = \Sigma_{i=1}^{p} N_i$ and (ℓ_i, m_i, n_i) be the direction cosines of the vector resultant of the sample observational vectors from the ith population and R_i the length of this resultant vector, then

$$\ell_i = \sum_{j=1}^{N_i} \ell_j / R_i$$

$$m_i = \sum_{j=1}^{N_i} m_j / R_i$$

$$n_i = \sum_{j=1}^{N_i} n_j \,/\, R_i \quad .$$

where the (ℓ_j, m_j, n_j) $j = 1,2,\ldots,N_i$ are the N_i vectors in the sample from the ith population so that

$$\cos \beta_j = \ell_j\lambda_i + m_j\mu_i + n_j\nu_i \quad .$$

Now for each i $(i = 1,2,\ldots,p)$ the logarithm of the likelihood, except for a constant term, is

$$\ell n \left(\prod_{j=1}^{N_i} \frac{\kappa}{\sinh \kappa} e^{\kappa \cos \beta_j} \right)$$

$$= N_i \,\ell n\, \kappa - N_i \,\ell n \sinh \kappa + \kappa \sum_{j=1}^{N_i} \cos \beta_j$$

$$= N_i(\ell n\, \kappa - \ell n \sinh \kappa) + \kappa \sum_{j=1}^{N_i} (\ell_j\lambda_i + m_j\mu_i + n_j\nu_i)$$

$$= N_i(\ell n\, \kappa - \ell n \sinh \kappa) + \kappa R_i(\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) \quad .$$

Thus for the entire set of observations the logarithm of the likelihood, except for a constant, is

$$(4.1.1) \quad N(\ell n\, \kappa - \ell n \sinh \kappa) + \kappa \sum_{i=1}^{p} R_i(\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) \quad .$$

Watson's test of coplanarity is essentially as follows: if a known direction (λ,μ,ν) is orthogonal to all the population means, then

the maximum likelihood (M.L.) estimator of (λ_i,μ_i,ν_i) is

$$(4.1.2) \qquad (\hat{\lambda}_i,\hat{\mu}_i,\hat{\nu}_i) = \left(\frac{(\ell_i-\lambda\cos\theta_i,\ m_i-\mu\cos\theta_i,\ n_i-\nu\cos\theta_i)}{\sin\theta_i}\right)$$

where

$$\cos\theta_i = \ell_i\lambda + m_i\mu + n_i\nu \quad .$$

Under these conditions the estimator of κ is given by the solution of

$$(4.1.3) \qquad \coth\hat{\kappa} - \frac{1}{\hat{\kappa}} = \frac{1}{N}\sum_{i=1}^{p} R_i \sin\theta_i \quad .$$

If no restrictions are placed on the mean vectors, then (ℓ_i,m_i,n_i) is the M.L. estimator of (λ_i,μ_i,ν_i) and the estimator of κ is the solution of

$$(4.1.4) \qquad \coth\kappa^* - \frac{1}{\kappa^*} = \frac{1}{N}\sum_{i=1}^{p} R_i \quad .$$

(The results (4.1.2), (4.1.3) and (4.1.4) will be verified in Section 4.2.)

Watson's method to test the null hypothesis that all the mean vectors lie in a plane normal to the prescribed direction (λ,μ,ν) requires Wilks' Theorem. It can be stated roughly as:

> When N is large, under the null hypothesis the distribution of $-2L$ is approximately a χ^2 distribution with degrees of freedom depending

on the number of parameters, where L is the logarithm of the likelihood-ratio.

(See Keeping [44], page 136). Thus Watson's method is the likelihood-ratio method.

In this case

$$L = \ell n \frac{g(\theta_i, \hat{\kappa})}{g(\theta_i, \kappa^*)}$$

$$= \ell n\, g(\theta_i, \hat{\kappa}) - \ell n\, g(\theta_i, \kappa^*)$$

where $\ell n\, g(\theta_i, \kappa)$ is the expression (4.1.1). Thus under the null hypothesis $\ell n\, g(\theta_i, \kappa)$ becomes

$$\ell n\, g(\theta_i, \hat{\kappa}) = N(\ell n\, \hat{\kappa} - \ell n \sinh \hat{\kappa}) + \hat{\kappa} \sum_{i=1}^{p} R_i \sin \theta_i \quad .$$

When κ and N are large ($\hat{\kappa}$ is large) the dispersion is small so that

$$\theta_i \simeq \pi/2$$

i.e.,

$$\sin \theta_i \simeq 1 \quad ,$$

and consequently

$$\sum_{i=1}^{p} R_i \sin \theta_i \simeq N \quad .$$

Thus

$$\ell n\, g(\theta_i, \hat{\kappa}) \simeq N(\ell n\, \hat{\kappa} - \ell n \sinh \hat{\kappa}) + \hat{\kappa} N$$

$$= N(\ell n\, \hat{\kappa} - \ell n \sinh \hat{\kappa} + \hat{\kappa})$$

$$= N(\ell n\, \hat{\kappa} - \ell n \tfrac{1}{2})$$

because for large $\hat{\kappa}$

$$\ell n \sinh \hat{\kappa} - \hat{\kappa} = \ell n \left(\frac{e^{\hat{\kappa}} - e^{-\hat{\kappa}}}{2e^{\hat{\kappa}}} \right) \simeq \ell n \frac{1}{2} \quad .$$

Similarly it can be shown that

$$\ell n\, g(\theta_i, \kappa^*) \simeq N(\ell n\, \kappa^* - \ell n \tfrac{1}{2})$$

Thus

$$L = N(\ell n\, \hat{\kappa} - \ell n \tfrac{1}{2}) - N(\ell n\, \kappa^* - \ell n \tfrac{1}{2})$$

$$= N(\ell n\, \hat{\kappa} - \ell n\, \kappa^*)$$

$$= N\, \ell n \left(\frac{\hat{\kappa}}{\kappa^*} \right) \quad .$$

and hence

(4.1.5) $$-2N\, \ell n \left(\frac{\hat{\kappa}}{\kappa^*} \right) \simeq \chi_p^2 \quad .$$

When κ is large we also have

$$\coth \kappa \simeq 1 \quad .$$

From (4.1.3) we obtain

(4.1.6) $$\hat{\kappa} \simeq \frac{N}{N - \sum_{i=1}^{p} R_i \sin \theta_i}$$

and from (4.1.4) we obtain

(4.1.7) $$\kappa^* \simeq \frac{N}{N - \sum_{i=1}^{p} R_i} .$$

Substituting (4.1.6) and (4.1.7) into (4.1.5) and simplifying we obtain

$$2N \ln \left[1 - \frac{\sum_{i=1}^{p} R_i (1 - \sin \theta_i)}{\sum_{i=1}^{p} (N_i - R_i)} \right] \simeq \chi_p^2 ,$$

which, when all N_i are large, reduces to

(4.1.8) $$\frac{2N \sum_{i=1}^{p} R_i (1 - \sin \theta_i)}{\sum_{i=1}^{p} (N_i - R_i)} \simeq \chi_p^2 .$$

Now if we let

$$\underline{\lambda} = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}$$

and

$$\underline{U} = \begin{pmatrix} \sum R_i \ell_i^2 & \sum R_i \ell_i m_i & \sum R_i \ell_i n_i \\ \sum R_i \ell_i m_i & \sum R_i m_i^2 & \sum R_i m_i n_i \\ \sum R_i \ell_i n_i & \sum R_i m_i n_i & \sum R_i n_i^2 \end{pmatrix}$$

then

$$\underline{\lambda}'\underline{U}\,\underline{\lambda} = \sum_{i=1}^{p} R_i(\ell_i\lambda + m_i\mu + n_i\nu)^2$$

$$= \sum_{i=1}^{p} R_i \cos^2 \theta_i \quad .$$

But if κ is large, the dispersion is small so that

$$\theta_i = \pi/2 + \epsilon_i$$

where ϵ_i is a very small number. Since

$$\cos^2 \theta_i = \epsilon_i^2 + O(\epsilon_i^4)$$

and

$$2(1 - \sin \theta_i) = \epsilon_i^2 + O(\epsilon_i^4)$$

then

$$2 \sum_{i=1}^{p} R_i(1 - \sin \theta_i) \simeq \sum_{i=1}^{p} R_i \cos^2 \theta_i \quad .$$

Thus (4.1.8) reduces to

$$(4.1.9) \qquad N \frac{\underline{\lambda}' \underline{U} \, \underline{\lambda}}{\sum_{i=1}^{p} (N_i - R_i)} \simeq \chi_p^2 \quad .$$

Hence for large N (4.1.9) provides an approximate test of the hypothesis that a set of p mean vectors lie in a plane normal to $\underline{\lambda}$.

The same likelihood-ratio procedure gives, as a test of coplanarity only, the statistic

$$(4.1.10) \qquad N \frac{\min_{\lambda} \underline{\lambda}' \underline{U} \, \underline{\lambda}}{\sum_{i=1}^{p} (N_i - R_i)} \simeq \chi_{p-2}^2$$

where the value of $\underline{\lambda}$ minimizing $\underline{\lambda}' \, \underline{U} \, \underline{\lambda}$ is the latent vector of $\underline{U}$ corresponding to its least latent root and this vector is the M.L. estimator of the normal to the best fitting plane.

4.2. The Generalization of Watson's Test of Coplanarity.

Using the same notation as in Section 4.1, we now consider the general approach to this test: suppose a known direction (λ, μ, ν) makes an angle α $(0 < \alpha < \pi)$ with all the population means $(\lambda_i, \mu_i, \nu_i)$.

The M.L. estimator of $(\lambda_i, \mu_i, \nu_i)$ is found by using Lagrangian multipliers. The logarithm of the likelihood of the entire set of observations (Equation (4.1.1)) becomes (except for a constant)

$$N(\ell n\ \kappa - \ell n \sinh \kappa) + \kappa \sum_{i=1}^{p} R_i(\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) \tag{4.2.1}$$
$$+ \sum_{i=1}^{p} \eta_i(\lambda_i^2 + \mu_i^2 + \nu_i^2 - 1) + \sum_{i=1}^{p} \xi_i(\lambda\lambda_i + \mu\mu_i + \nu\nu_i - \cos\alpha)$$

where the constraints are

$$\lambda\lambda_i + \mu\mu_i + \nu\nu_i = \cos\alpha \ , \tag{4.2.2}$$
$$\lambda_i^2 + \mu_i^2 + \nu_i^2 = 1$$

and $\{\eta_i\}$ and $\{\xi_i\}$ are Lagrange multipliers.

By taking the partial derivatives of (4.2.1) with respect to λ_i, μ_i and ν_i (for some $i = 1,\ldots,p$) and equating the result to zero we obtain

$$\kappa R_i \ell_i + 2\eta_i \lambda_i + \xi_i\lambda = 0 \tag{4.2.3}$$

$$\kappa R_i m_i + 2\eta_i \mu_i + \xi_i\mu = 0 \tag{4.2.4}$$

$$\kappa R_i n_i + 2\eta_i \nu_i + \xi_i\nu = 0 \ . \tag{4.2.5}$$

Multiplying (4.2.3) by λ_i, (4.2.4) by μ_i and (4.2.5) by ν_i and adding the three equations we obtain

$$\kappa R_i (\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) + 2\eta_i + \xi_i \cos\alpha = 0 \ . \tag{4.2.6}$$

Multiplying (4.2.3) by λ, (4.2.4) by μ, and (4.2.5) by ν, and

adding we obtain

(4.2.7) $$\kappa R_i (\ell_i\lambda + m_i\mu + n_i\nu) + 2\eta_i \cos\alpha + \xi_i = 0 \quad .$$

From (4.2.6) and (4.2.7) by elimination

(4.2.8) $$\xi_i = \frac{\kappa R_i [\cos\alpha (\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) - (\ell_i\lambda + m_i\mu + n_i\nu)]}{\sin^2\alpha}$$

and

(4.2.9) $$2\eta_i = \frac{\kappa R_i [\cos\alpha (\ell_i\lambda + m_i\mu + n_i\nu) - (\ell_i\lambda_i + m_i\mu_i + n_i\nu_i)]}{\sin^2\alpha} \quad .$$

Multiplying (4.2.3) by ℓ_i , (4.2.4) by m_i , and (4.2.5) by n_i and adding we obtain

(4.2.10) $$\kappa R_i + 2\eta_i (\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) + \xi_i(\ell_i\lambda + m_i\mu + n_i\nu) = 0 \quad .$$

Using (4.2.8) and (4.2.9) this reduces to

$$\sin^2\alpha + 2\cos\alpha (\ell_i\lambda + m_i\mu + n_i\nu)(\ell_i\lambda_i + m_i\mu_i + n_i\nu_i)$$
$$- (\ell_i\lambda_i + m_i\mu_i + n_i\nu_i)^2 - (\ell_i\lambda + m_i\mu + n_i\nu)^2 = 0$$

(provided $\kappa \neq 0$ and $R_i \neq 0$), which is of the form

$$[\ell_i\lambda_i + m_i\mu_i + n_i\nu_i]^2 + B(\ell_i\lambda_i + m_i\mu_i + n_i\nu_i) + C = 0$$

where $B = -2\cos\alpha (\ell_i\lambda + m_i\mu + n_i\nu)$

$C = (\ell_i\lambda + m_i\mu + n_i\nu)^2 - \sin^2\alpha$.

Since this is a quadratic equation, we have

$$\ell_i\lambda_i + m_i\mu_i + n_i\nu_i \tag{4.2.11}$$

$$= \cos\alpha\,(\ell_i\lambda + m_i\mu + n_i\nu) \pm \sin\alpha\,[1 - (\ell_i\lambda + m_i\mu + n_i\nu)^2]^{1/2} \quad .$$

For convenience, let

$$\cos\beta_i = \ell_i\lambda_i + m_i\mu_i + n_i\nu_i \tag{4.2.12}$$

$$\cos\theta_i = \ell_i\lambda + m_i\mu + n_i\nu \quad . \tag{4.2.13}$$

Then (4.2.11) can be written as

$$\cos\beta_i = \cos\alpha\cos\theta_i \pm \sin\alpha(1 - \cos^2\theta_i)^{1/2}$$

$$= \cos\alpha\cos\theta_i \pm \sin\alpha\sin\theta_i$$

$$= \cos(\alpha \mp \theta_i) = \cos(\theta_i \mp \alpha) \quad .$$

Since

$$0 < \alpha\ ,\ \beta_i\ ,\ \theta_i < \pi$$

then

$$\beta_i = \alpha \mp \theta_i \quad .$$

From the geometry of the sphere it follows

$$\beta_i = \alpha - \theta_i \quad . \tag{4.2.14}$$

Now (4.2.8) becomes

$$\xi_i = \kappa R_i [\cos\alpha \cos\beta_i - \cos\theta_i] / \sin^2\alpha$$

$$= \kappa R_i [\cos\alpha \cos(\alpha - \theta_i) - \cos\theta_i] / \sin^2\alpha$$

$$= \kappa R_i [\cos\alpha (\cos\alpha \cos\theta_i + \sin\alpha \sin\theta_i) - \cos\theta_i] / \sin^2\alpha$$

$$= \kappa R_i [\cos\alpha \sin\theta_i - \sin\alpha \cos\theta_i] / \sin\alpha \quad ,$$

i.e.,

(4.2.15) $$\xi_i = \kappa R_i \sin(\theta_i - \alpha) / \sin\alpha \quad ,$$

and (4.2.9) becomes

$$2\eta_i = \kappa R_i [\cos\alpha \cos\theta_i - \cos\beta_i] / \sin^2\alpha$$

$$= \frac{\kappa R_i}{\sin^2\alpha} [\cos\alpha \cos\theta_i - \cos\alpha \cos\theta_i - \sin\alpha \sin\theta_i] \quad ,$$

i.e.,

(4.2.16) $$2\eta_i = -\kappa R_i \frac{\sin\theta_i}{\sin\alpha} \quad ,$$

and substituting these results into (4.2.3) we obtain

(4.2.17) $$\hat{\lambda}_i = \frac{\ell_i \sin\alpha + \lambda \sin(\theta_i - \alpha)}{\sin\theta_i} \quad .$$

In a similar manner we also obtain

(4.2.18) $$\hat{\mu}_i = \frac{m_i \sin\alpha + \mu \sin(\theta_i - \alpha)}{\sin\theta_i}$$

$$\hat{\nu}_i = \frac{n_i \sin \alpha + \nu \sin (\theta_i - \alpha)}{\sin \theta_i} \quad . \tag{4.2.19}$$

The results (4.2.17), (4.2.18) and (4.2.19) define the M.L. estimator of $(\lambda_i, \mu_i, \nu_i)$ under the null hypothesis that the known (λ, μ, ν) makes an angle α with all the population means.

Under this null hypothesis the M.L. estimator of κ can be found by differentiating (4.2.1) with respect to κ and equating the result to zero. The equation that the M.L. estimator of κ must satisfy is

$$\coth \hat{\kappa} - \frac{1}{\hat{\kappa}} = \frac{1}{N} \sum_{i=1}^{p} R_i \cos \beta_i = \frac{1}{N} \sum_{i=1}^{p} R_i \cos (\alpha - \theta_i) \quad . \tag{4.2.20}$$

When no restrictions are placed on the mean vectors, the resultant vector (ℓ_i, m_i, n_i) is the M.L. estimator of $(\lambda_i, \mu_i, \nu_i)$. Thus under the alternative hypothesis, the equation that the M.L. estimator of κ must satisfy is

$$\begin{aligned} \coth \kappa^* - \frac{1}{\kappa^*} &= \frac{1}{N} \sum_{i=1}^{p} R_i (\ell_i \ell_i + m_i m_i + n_i n_i) \\ &= \frac{1}{N} \sum_{i=1}^{p} R_i \quad . \end{aligned} \tag{4.2.21}$$

We now proceed in the same way as in Section 4.1. The logarithm of the likelihood-ratio can be written as

$$L = \ell n \; g_\alpha (\theta_i, \hat{\kappa}) - \ell n \; g_\alpha (\theta_i, \kappa^*)$$

where $\ell n\, g_\alpha\,(\theta_i,\kappa)$ is the expression (4.1.1) considered in relation to general α (hence the subscript α). Under the null hypothesis $\ell n\, g_\alpha\,(\theta_i,\kappa)$ becomes

$$(4.2.22)\qquad \ell n\, g_\alpha\,(\theta_i,\hat{\kappa})$$

$$= N(\ell n\,\hat{\kappa} - \ell n \sinh \hat{\kappa}) + \hat{\kappa} \sum_{i=1}^{p} R_i(\ell_i\hat{\lambda}_i + m_i\hat{\mu}_i + n_i\hat{\nu}_i)$$

where by (4.2.17), (4.2.18) and (4.2.19),

$$\ell_i\hat{\lambda}_i + m_i\hat{\mu}_i + n_i\hat{\nu}_i$$

$$= \frac{[\sin\alpha + (\ell_i\lambda + m_i\mu + n_i\nu)\sin(\theta_i - \alpha)]}{\sin\theta_i}$$

$$= \frac{\sin\alpha + \cos\theta_i\,(\sin\theta_i\cos\alpha - \cos\theta_i\sin\alpha)}{\sin\theta_i}$$

$$= \sin\alpha\sin\theta_i + \cos\alpha\cos\theta_i$$

$$= \cos\,(\theta_i - \alpha)\quad ,$$

so that (4.2.22) becomes

$$(4.2.23)\qquad \ell n\, g_\alpha(\theta_i,\hat{\kappa}) = N(\ell n\,\hat{\kappa} - \ell n \sinh \hat{\kappa}) + \hat{\kappa} \sum_{i=1}^{p} R_i \cos\,(\theta_i - \alpha)\quad .$$

When κ and N are large the dispersion is small so that

$$\theta_i \simeq \alpha$$

and consequently

$$\sum_{i=1}^{p} R_i \cos(\theta_i - \alpha) \simeq N \ .$$

Thus

$$(4.2.24) \qquad \ln g_\alpha(\theta_i, \hat{\kappa}) \simeq N(\ln \hat{\kappa} - \ln \sinh \hat{\kappa}) + \hat{\kappa} N$$

$$\simeq N(\ln \hat{\kappa} - \ln \tfrac{1}{2})$$

because $\ln \sinh \hat{\kappa} - \hat{\kappa} \simeq \ln \frac{1}{2}$ for large $\hat{\kappa}$. Similarly it follows

$$(4.2.25) \qquad \ln g_\alpha(\theta_i, \overset{*}{\kappa}) \simeq N(\ln \kappa^* - \ln \tfrac{1}{2}) \ .$$

Thus we have

$$(4.2.26) \qquad L = N \ln\left(\frac{\hat{\kappa}}{\kappa^*}\right) ,$$

and by Wilk's theorem we obtain

$$(4.2.27) \qquad -2N \ln\left(\frac{\hat{\kappa}}{\kappa^*}\right) \simeq \chi_p^2 \ .$$

When κ is large we also have

$$\coth \kappa \simeq 1 \ .$$

Then by (4.2.20) the approximate estimate of κ under the null hypothesis is

$$(4.2.28) \qquad \hat{\kappa} \simeq \frac{N}{N - \sum_{i=1}^{p} R_i \cos(\alpha - \theta_i)} ,$$

and by (4.2.21) the approximate estimate of κ under the alternative hypothesis is

(4.2.29) $$\kappa^{*} \simeq \frac{N}{N - \sum_{i=1}^{p} R_i} .$$

Substituting (4.2.28) and (4.2.29) into the left hand side of (4.2.27) we obtain

$$- 2 N \ell n \left(\frac{\hat{\kappa}}{\kappa^{*}} \right)$$

$$= 2 N \ell n \left(\frac{\kappa^{*}}{\hat{\kappa}} \right)$$

$$= 2 N \ell n \left(\frac{\sum_{i=1}^{p} [N_i - R_i + R_i - R_i \cos (\alpha - \theta_i)]}{\sum_{i=1}^{p} (N_i - R_i)} \right)$$

$$= 2 N \ell n \left(1 + \frac{\sum_{i=1}^{p} R_i [1 - \cos (\alpha - \theta_i)]}{\sum_{i=1}^{p} (N_i - R_i)} \right)$$

and when all the N_i are large

$$\frac{\dfrac{\sum_{i=1}^{p} R_i[1 - \cos(\alpha - \theta_i)]}{\sum_{i=1}^{p} (N_i - R_i)}}{\ell n\left(1 + \dfrac{\sum_{i=1}^{p} R_i[1 - \cos(\alpha - \theta_i)]}{\sum_{i=1}^{p} (N_i - R_i)}\right)} \simeq 1 ,$$

so that (4.2.27) reduces to

$$(4.2.30) \qquad 2N \frac{\sum_{i=1}^{p} R_i[1 - \cos(\alpha - \theta_i)]}{\sum_{i=1}^{p} (N_i - R_i)} \simeq \chi_p^2 ,$$

where α and θ_i are defined in (4.2.2) and (4.2.13) respectively. The statistic (4.2.30) can be used as a significance test of the null hypothesis that the population mean vectors make an angle α $(0 < \alpha < \pi)$ with a prescribed direction (λ, μ, ν), provided κ is large.

We note here that if $\alpha = \pi/2$ the statistic (4.2.30) reduces to

$$2N \frac{\sum_{i=1}^{p} R_i (1 - \sin\theta_i)}{\sum_{i=1}^{p} (N_i - R_i)}$$

which is (4.1.8), demonstrating that Watson's statistic is indeed the particular case of $\alpha = \pi/2$.

The cases $\kappa = 0$ and $R_i = 0$ are degenerate trivial cases.

Another approximate form can be obtained when κ is large. In this case the dispersion is small so that we can let

$$\theta_i = \alpha - \epsilon_i \quad .$$

Then

$$2(1 - \cos (\alpha - \theta_i)) = 2(1 - \cos \epsilon_i)$$
$$= \epsilon_i^2 + O(\epsilon_i^4)$$

and

$$\sin^2 (\alpha - \theta_i) = \epsilon_i^2 + O(\epsilon_i^4) \quad ,$$

so that (4.2.30) has the further approximate form

$$N \frac{\sum_{i=1}^{p} R_i \sin^2 (\alpha - \theta_i)}{\sum_{i=1}^{p} (N_i - R_i)} \simeq \chi_p^2 \quad . \tag{4.2.31}$$

In conclusion, either (4.2.30) or (4.2.31) provide a test statistic for the hypothesis that the ends of the mean vectors of p populations all lie on a circular section of the unit sphere when the populations are distributed according to the Fisher distribution.

4.3 Remarks on Further Study

To obtain an analogous statistic of (4.1.9) we must minimize (4.2.30) with respect to λ, μ, ν and α subject to $\lambda^2 + \mu^2 + \nu^2 = 1$. This is equivalent to minimizing the expression

$$\sum_{i=1}^{p} R_i[1 - \cos(\alpha - \theta_i)] \text{ (subject to } \lambda^2 + \mu^2 + \nu^2 = 1)$$

with respect to these four parameters. This has been done by restricting α $(0 < \alpha < \pi)$ to $\alpha \neq \pi/2$ and obtaining

$$\tan \hat{\alpha} = \frac{\sum_{i=1}^{p} R_i \sin \theta_i}{\sum_{i=1}^{p} R_i \cos \theta_i}$$

and a system of three equations that the minimizing values of λ, μ and ν would have to satisfy. However, the form of this system does not allow for further reduction. This problem at present is an open question.

The statistic (4.2.30) is useful for the case of known α. The problem of obtaining a test statistic for unknown α could also be investigated.

The statistic (4.2.30) is an approximation because N (as well as each N_i) and κ were assumed to be large. Without these assumptions it seems possible that an exact test could be derived. For small κ, the estimation equations for κ in the case of $\alpha = \pi/2$ (as well as in the general case) are difficult to reduce to a suitable workable form.

Only a very complicated form of L is possible and a satisfactory form has not yet been derived.

For the case $\alpha = \pi/2$ an analysis of variance analogue for the statistic (4.1.9) has been suggested by Watson [96]. It has the form

$$\text{(4.3.1)} \qquad \frac{N-p}{p} \; \frac{\underline{\lambda}' \, \underline{U} \, \underline{\lambda}}{\sum_{i=1}^{p} (N_i - R_i)} \simeq F_{p,2(N-p)} \quad .$$

For large N this is the same as (4.1.9). Thus (4.1.9) and (4.3.1) would provide two approximate tests of the hypothesis that a set of p mean vectors lie in a plane normal to $\underline{\lambda}$. Watson has also shown by a numerical example that (4.3.1) could be conveniently arranged in an analysis of variance table. The problem of deriving an analysis of variance analogue for the generalized test statistic (4.2.30) should also be pursued.

BIBLIOGRAPHY

[1] Arnold, K. J., "On Spherical Probability Distributions". Ph. D. Dissertation, 1941, Massachusetts Institute of Technology.

[2] Bachelier, L., "Théorie de la spéculation". Annales Scientifiques de l'école Normale Supérieure, 1900, 17, 21-88.

[3] Bennett, W. R., "Distribution of the Sum of Randomly Phased Components". The Quarterly of Applied Mathematics, 1947, 5, 385-393.

[4] Bernoulli, D., "Recherches Physiques et Astronomiques sur le Problème proposé pour la Seconde Fois par l'Académie Royale des Sciences de Paris". l'Académie Royale des Sciences, 1734, Tome III, 95-134.

[5] Bingham, C., "Distributions on the Sphere and on the Projective Plane". Ph. D. Dissertation, 1964, Yale University.

[6] Breitenberger, E., "Theory of Multiple Scattering". Proceedings of the Royal Society, Series A, 1959, 250, 514-523.

[7] Breitenberger, E., "Analogues of the Normal Distribution on the Circle and the Sphere". Biometrika, 1963, 50, 81-88.

[8] Brooks, C. E. P. and Carruthers, N., "Handbook of Statistical Methods in Meteorology". H. M. Stationary Office, 1953, London, England.

[9] Bruckshaw, J. McG., and Vincenz, S. A., "The Permanent Magnetism of the Mull Lavas". Monthly Notices of the Royal Astronomical Society, Geophysical Supplement, 1954, 6, 579-589.

[10] Chandrasekhar, S., "Stochastic Problems in Physics and Astronomy". Reviews of Modern Physics, 1943, 15, 1-89.

[11] Coulson, C. A., "Note on the Random-Walk Problem". Proceedings of the Cambridge Philosophical Society, 1947, 43, 583-586.

[12] Cramér, H., "Random Variables and Probability Distributions". Cambridge University Press, 1937, Cambridge, England.

[13] Daniels, H. E., "The Covering Circle of a Sample from a Circular Normal Distribution". Biometrika, 1952, 39, 137-143.

[14] Downs, T. D., "The von Mises Distribution: Derived Distributions, Regression Theory, and Some Applications to Meteorological Problems". Ph. D. Dissertation, 1965, University of Michigan.

[15] Downs, T. D., "Some Relationships Among the von Mises Distributions of Different Dimensions". Biometrika, 1966, 53, 269-272.

[16] Durand, D. and Greenwood, J. A., "The Integral Solution of Pearson's Random Walk Problem and Related Matters". The Annals of Mathematical Statistics, 1953, 24, 686-687.

[17] Durand, D. and Greenwood, J. A., "Random Unit Vectors II: Usefulness of Gram-Charlier and Related Series in Approximating Distributions". The Annals of Mathematical Statistics, 1957, 28, 978-986.

[18] Durand, D. and Greenwood, J. A., "Modifications of the Rayleigh Test for Uniformity in Analysis of Two-Dimensional Orientation Data". The Journal of Geology, 1958, 66, 229-238.

[19] Einstein, A., "Über die von der Molekularkinetischen Theorie der Wärme geforderte Bewegung von in ruhenden Flüssigkeiten suspendierten Teilchen". Annalen der Physik, Leipzig, Germany, 1905, 17, 549-560.

[20] Epstein, B. and Sobel, M., "Life Testing". The Journal of the American Statistical Association, 1953, 48, 486-502.

[21] Erdélyi, A., "Higher Transcendental Functions". Bateman Manuscript Project, 1953, Volumes I and II. (See Vol. II, page 355), McGraw-Hill, New York.

[22] Esseen, C. G., "Fourier-Analysis of Distribution Functions". Acta Mathematica, 1945, 77, 1-125.

[23] Fairbairn, H. W., "Structural Petrology of Deformed Rocks". Addison-Wesley, Cambridge, 1949.

[24] Favro, L. D., "Theory of the Rotational Brownian Motion of a Free Rigid Body". Physics Review, second series, 1960, 119, Part I, 53-62.

[25] Fessenden, R. F., "The Problem of the Random Path". Nature, 1907, 75, 392.

[26] Fisher, R. A., "Tests of Significance in Harmonic Analysis". The Proceedings of the Royal Society, Series A, 1929, 125, 54-59.

[27] Fisher, R. A., "Dispersion on a Sphere". Proceedings of the Royal Society, London, Series A, 1953, 217, 295-305.

[28] Furry, W. H., "Isotropic Rotational Brownian Motion". Physics Review, second series, 1957, 107, 7-12.

[29] Gauss, K. F., "Theoria Motus Corporum Coelestium". Hamburg, Germany, 1809, Section 177.

[30] Greenwood, J. A., "Distribution Theory of Some Angular Variables". Ph. D. Dissertation, 1959, Harvard University.

[31] Greenwood, J. A. and Durand, D., "The Distribution of Length of Components of the Sum of n Random Unit Vectors". The Annals of Mathematical Statistics, 1955, 26, 233-246.

[32] Gumbel, E. J., "Applications of the Circular Normal Distribution". The Journal of the American Statistical Association, 1954, 49, 267-297.

[33] Gumbel, E. J., Greenwood, J. A. and Durand, D., "The Circular Normal Distribution: Theory and Tables". The Journal of the American Statistical Association, 1953, 48, 131-152.

[34] deHaas-Lorentz, G. L., "Die Brownsche Bewegung und einige verwandte Erscheinungen". Brunswick: Friedrick Vieweg und Sohn, 1913.

[35] Hall, P., "The Distribution of Means for Samples of Size N Drawn from a Population in which the Variate Take Values Between 0 and 1 , All Such Values Being Equally Probable". Biometrika, 1927, 19, 240-245.

[36] Hartman, P. and Wintner, A., "On the Spherical Approach to the Normal Distribution Law". The American Journal of Mathematics, 1940, 62, 759-779.

[37] Hobson, E. W., "The Theory of Spherical and Ellipsoidal Harmonics". Cambridge University Press, 1931, see Chapter 2.

[38] Horner, F., "A Problem on the Summation of Simple Harmonic Functions of the Same Amplitude and Frequency but of Random Phase". The Philosophical Magazine, Series 7, 1946, 37, 145-162.

[39] Irwin, J. O., "On the Frequency Distribution of the Means of Samples from a Population Having any Law of Frequency with Finite Moments, with Special Reference to Pearson's Type II". Biometrika, 1927, 19, 225-239.

[40] Jeffreys, H., "The Theory of Probability". Oxford University Press, Oxford, England, 2nd Edition, 1948, see page 302.

[41] Joos, G., "Theoretical Physics". Blackie and Son Limited, London and Glasgow, England, 1947.

[42] Kac, M., "On a Characterization of the Normal Distribution". The American Journal of Mathematics, 1939, 61, 726-728.

[43] Kac, M. and van Kampen, E. R., "Circular Equidistributions and Statistical Independence". The American Journal of Mathematics, 1939, 61, 677-682.

[44] Keeping, E. S., "An Introduction to Statistical Inference". D. Van Nostrand Company, Incorporated, Princeton, New Jersey, 1962, see page 136.

[45] Kluyver, J. C., "A Local Probability Problem - Random Walk". Koninklijke Akademie van Wetenshappen - Amsterdam Proceedings of the Section of Sciences Vol. 8, Pt. 1, 1906, pp. 341-350.

[46] Krumbein, W. C., "Preferred Orientation of Pebbles in Sedimentary Deposits". The Journal of Geology, 1939, 47, 673-706.

[47] Laplace, P. S., "Théorie analytique des Probabilitiés". Paris, 1812, livre II, Chapter IV (Following a Memoir of 1811).

[48] Legendre, A. M., "Nouvelles Méthodes pour la Détermination des Orbites des Comètes". Paris, 1806, see page 72.

[49] Lévy, P., "Sur l'Arithmétique des Lois de Probabilités Enroulées". Comptes Rendus, Bulletin de la Société Mathématique de France, 1938, 32-34.

[50] Lévy, P., "Observation sur la Communication Précédente". Comptes Rendus, Bulletin de la Société Mathématique de France, 1938, 36-37.

[51] Lévy, P., "l'Addition des Variables Aléatoires, définies sur une Circonférence". Comptes Rendus, Bulletin de la Société Mathématique de France, 1939, 67, 1-41.

[52] Lord, R. D., "A Problem of Random Vectors". Philosophical Magazine, 1948, 39, 66-71.

[53] Lord, R. D., "The Use of the Hankel Transform in Statistics. I. General Theory and Examples." Biometrika, 1954, 41, 44-55.

[54] Lord, R. D., "The Use of the Hankel Transform in Statistics. II. Methods of Computation." Biometrika, 1954, 41, 344-350.

[55] Marcinkiewicz, J., "Sur les Variables Aléatoires Enroulées". Comptes Rendus, Bulletin de la Société Mathématique de France, 1938, 34-36.

[56] Mises, R. von, "Über die 'Ganzzahligkeit' der Atomgewichte und verwandte Fragen". Physikalische Zeitschrift, 1918, 19, 490-500.

[57] Nicholson, J. W., "Generalisation of a Theorem due to Sonine". The Quarterly Journal of Pure and Applied Mathematics, 1920, 48, 321-329.

[58] Pearson, K., "The Problem of the Random Walk". Nature, 1905, 72, 294.

[59] Pearson, K., "The Problem of the Random Walk". Nature, 1905, 72, 318.

[60] Pearson, K., "On a Mathematical Theory of Random Migration". Drapers' Company Research Memoirs, Biometric Series, Number 3, 1906.

[61] Perrin, F., "Étude Mathématique du Mouvement Brownien de Rotation". Annales Scientifiques de l'école Normale Supérieure, 1928, (3), 45, 1-51.

[62] Perrin, F., "Mouvement Brownien d'un Ellipsoïde". The Journal of Physical Radium, 1934, 5, 497-511.

[63] Pincus, H. J., "The Analysis of Aggregates of Orientation Data in the Earth Sciences". The Journal of Geology, 1953, 61, 482-509.

[64] Pólya, G., "Elementarer Beweis einer Thetaformel". S. B. Akad. Wiss. Berlin. 1927. See pages 158-161.

[65] Pólya, G., "Sur Quelques Points de la Théorie des Probabilités". Paris, Université, Faculté des Sciences, Annales de Institu de Henri Poincaré, 1931, 1, 117-161.

[66] Quenouille, M. H., "On the Problem of Random Flights". Proceedings of the Cambridge Philosophical Society, 1947, 43, 581-582.

[67] Rayleigh, Lord (J. W. Strutt), "On the Resultant of a large Number of Vibrations of the same Pitch and of arbitrary Phase". Philosophical Magazine and Journal of Science, Series 5, 1880, 10, No. 60, 73-78.

[68] Rayleigh, Lord (J. W. Strutt), "The Problem of the Random Walk". Nature, 1905, 72, 318.

[69] Rayleigh, Lord (J. W. Strutt), "On the Problem of Random Vibrations and of Random Flights in One, Two, or Three Dimensions". Philosophical Magazine and Journal of Science, Series 6, 1919, No. 220, 37, 321-347.

[70] Rayleigh, Lord (J. W. Strutt), "On the Resultant of a Number of unit Vibrations, whose Phases are at random over a Range not limited to an Integral Number of Periods". Philosophical Magazine and Journal of Science, Series 6, 1919, No. 220, 37, 498-515.

[71] Roberts, P. H. and Ursell, H. D., "Random Walk on a Sphere and on a Riemannian Manifold". Philosophical Transactions, Series A, 1960, 252, 317-356.

[72] Ruben, H., "Probability Content of Regions under Spherical Normal Distributions, I ". The Annals of Mathematical Statistics, 1960, 31, 598-618.

[73] Ruben, H., "Probability Content of Regions under Spherical Normal Distributions, II : The Distribution of the Range in Normal Samples". The Annals of Mathematical Statistics, 1960, 31, 1113-1121.

[74] Runcorn, S. K., "Rock Magnetism". Science, 1959, 129, 1002-1012.

[75] Schoenberg, I. J., "Metric Spaces and Completely Monotone Functions". The Annals of Mathematics, 1938, 39, 811-841.

[76] Selby, B., "Girdle Distributions on a Sphere". Biometrika, 1964, 51, 381-392.

[77] Slack, M., "The Probability Distributions of Sinusoidal Oscillations Combined in Random Phase". The Journal of the Institute of Electrical Engineering, 1946, Part 3, 93, 76-96.

[78] Sommerfeld, A., "Vorlesungen über Theoretische Physik". Leipzig: Akademische Verlagsges, 1947, Volume 7. (See also "Partial Differential Equations in Physics". New York: Academic Press, 1949.)

[79] Sonine, N. J., Mathematical Annals, 1880, Volume 16, (See page 41 and 46).

[80] Spitzer, F., "Principles of Random Walk". D. Van Nostrand Company, 1964, Princeton, New Jersey.

[81] Stephens, M. A., "The Statistics of Directions: the von Mises and Fisher Distributions". Ph. D. Dissertation, 1962, University of Toronto.

[82] Stephens, M. A., "Exact and Approximate Test for Directions". Biometrika, 1962, 49, 463-477.

[83] Stephens, M. A., "Exact and Approximate Tests for Directions II ". Bimetrika, 1962, 49, 547-552.

[84] Stephens, M. A., "The Distribution of the Goodness-of-fit Statistic U_n^2 , I ". Biometrika, 1963, 50, 303-313.

[85] Stephens, M. A., "Random Walk on a Circle". Biometrika, 1963, 50, 385-390.

[86] Stephens, M. A., "The Testing of Unit Vectors for Randomness". The Journal of the American Statistical Association, 1964, 59, 160-167.

[87] Stephens, M. A., "Percentage Points of Statistics for Testing Randomness of Directions". Office of Naval Research, John Hopkins University, Technical Report No. 27, June 1965.

[88] Stephens, M. A., "Tests for the Dispersions, and for the Modal Vector of a Unimodal Distribution on a Sphere". Office of Naval Research, John Hopkins University, Technical Report No. 31, August 1965.

[89] Stephens, M. A., "Statistics Connected with the Uniform Distribution: Percentage Points and Application to Testing for Randomness of Directions". Biometrika, 1966, 53, 235-240.

[90] 'Student' (W. S. Gosset), "The Probable Error of a Mean". Biometrika, 1908, 6, 1-25.

[91] Tucker, G. B., "Upper Winds over the World, Part III, Standard Vector Deviations of Wind up to the 100 Millebar Level over the World". Great Britain Meteorological Office, Geophysical Memoirs, 1961, No. 105, Volume 13.

[92] Waterman, T. H., "The Analysis of Spatial Orientation". Ergebnisse der Biologie, 1963, 26, 98-117.

[93] Watson, G. N., "A Treatise on the Theory of Bessel Functions". Cambridge University Press, 1922, (See 2nd Edition, 1944).

[94] Watson, G. S., "Analysis of Dispersion on a Sphere". Monthly Notices of the Royal Astronomical Society, Geophysical Supplement, 1956, 7, 153-159.

[95] Watson, G. S., "A Test for Randomness of Directions". Monthly Notices of the Royal Astronomical Society, Geophysical Supplement, 1956, 7, 160-161.

[96] Watson, G. S., "More Significance Tests on the Sphere". Biometrika, 1960, 47, 87-91.

[97] Watson, G. S., "Goodness-of-Fit Tests on a Circle I ". Biometrika 1961, 48, 109-114.

[98] Watson, G. S., "Goodness-of-Fit Tests on a Circle II ". Biometrika, 1962, 49, 57-63.

[99] Watson, G. S., "Equatorial Distributions on a Sphere". Biometrika, 1965, 52, 193-201.

[100] Watson, G. S. and Irving, E., "Statistical Methods in Rock Magnetism". Monthly Notices of the Royal Astronomical Society, 1957, 7, 289-300.

[101] Watson, G. S. and Williams, E. J., "On the Construction of Significance Tests on the Circle and the Sphere". Biometrika, 1956, 43, 344-352.

[102] Whittaker, E. T. and Robinson, G., "The Calculus of Observations". Blackie and Son Limited, London, England, 1924, (See 4th Edition, 1944, 209-220).

[103] Wiener, N., "The Average of an Analytic Functional and the Brownian Movement". The Proceedings of the National Academy of Science, U. S. A., 1921, 7, 294-298.

[104] Wintner, A., "On the Stable Distribution Laws". The American Journal of Mathematics, 1933, 55, 335-339.

[105] Wintner, A., "On a Class of Fourier Transforms". The American Journal of Mathematics, 1936, 58, 45-90.

[106] Wintner, A., "On the Shape of the Angular Case of Cauchy's Distribution Curves". The Annals of Mathematical Statistics, 1947, 18, 589-593.

[107] Zernike, F., "Wahrscheinlichkeitsrechnung und Mathematische Statistik". Handbuch der Physik, Berlin, Julius Springer, 1928, 3, 419-492, (See 477-478).